N73-31348
CR-133982

Phase 1 of the Earth
Resources Data Analysis
Program

FINAL REPORT
(June 1972 - May 1973)

ICSA

INSTITUTE FOR COMPUTER SERVICES AND APPLICATIONS

RICE UNIVERSITY

Phase 1 of the Earth
Resources Data Analysis
Program

CA

FINAL REPORT

(June 1972 - May 1973)

Institute for Computer Services and Applications
Rice University
Houston, Texas 77001

June, 1973

This research was sponsored by NASA
under Contract NAS 9-12776

I. INTRODUCTION

Phase I of the Earth Resources Data Analysis Program at Rice University has been marked by many diverse efforts. In addition to the various projects undertaken, several tools have been developed. These include the installation of LARSYS on the ICSA's IBM 370/155 (see Appendix A) and the establishment of a framework at Rice for both performing the various tasks and providing an effective communications channel with NASA JSC personnel.

Though most projects are currently at some stopping point, others will require further study. In addition, further analysis is being conducted in many of these areas to supplement existing conclusions. However, all projects will issue reports by the end of this summer.

It should be mentioned that, like most research efforts, some research efforts led to "blind alleys". In particular, solutions to the optimal feature extraction problem in generality proved elusive; efforts were redirected to finding solutions that minimized a bound on the probability of misclassification. On the other hand, new ideas not originally anticipated at the start of the contract have appeared such as use of the Cholesky decomposition and use of spatial information.

Each of the projects undertaken is briefly discussed in the next section. More details may be found in the accompanying reports. (see Appendix B.)

II. AREAS STUDIED

A variety of problems were considered. Some projects have been completed, while others are currently unfinished. However most of these projects have resulted in indicating further areas to be studied. Recommendation for future study are included herein.

The projects undertaken include:

- Use of the Cholesky decomposition in feature selection and classification algorithms
- Optimal feature selection and extraction
- Probability density estimation and non-parametric classifiers
- Use of spatial information in classification

—— Model for crop row reflectance
—— Miscellaneous projects

Technical reports presently available on these subject are included Appendix B. Other reports are forthcomingasindicated below.

A. Use of the Cholesky decomposition in divergence and classification algorithms

The Cholesky decomposition is the factorization of a symmetric, positive-definite matrix into a product of a lower triangular matrix and its transpose, i.e.

$$K = LL^{*}$$

or (the modified Cholesky decomposition), where D is diagonal.

$$K = L^{1}D^{1}L^{1*}$$

The covariance matrices are symmetric, positive-definite, and therefore, these decompositions may be applied to them. With these decompositions, expressions involving covariance matrices may be simplified and computations reduced.

We derived the simplified expressions (see Report 8) and tested programs using this formulation, using LARSYS as a framework. In the divergence calculation, a time savings of a factor of 2-4 is effected, and in the maximum likelihood classifier the corresponding factor is ~30% . Also the resultant computations are more numerically stable. Thus these algorithms represent a significant improvement over existing routines and should be included in all present systems (e.g. ERIPS) where such calculations are performed. Future work in this area will include a study of how to treat ill-conditioned systems.

B. Optimal feature selection and feature extraction

Early efforts to find explicit solutions to this problem led to apparently insoluble systems of equations. Efforts were then directed to obtaining as optimal solutions as possible. This led to a formulation that minimizes a bound on the probability of misclassification. Since the equations are still quite difficult to solve and one is not sure how "tight" the bound is, it was advisable to search for a simpler special case to treat.

The special case where only two pattern classes are involved is more amenable to a solution. Here the bound on the probability of misclassification is obtained by using the Bhattacharyya distance. Two special subcases appear that are relatively easily soluble: case (i) where the two covariance matrices are equal but the means unequal, and case (ii) where the means are equal but not the covariance matrices.

These two cases were programmed and tested using Purdue Flight Line C1 data. Though the exact conditions of these cases could not be met with this data, close approximations were used for testing purposes. Classification results using the generated feature extractors were compared with those using the best single channel selected by the divergence criterion. Results were encouraging, in that cases more closely fitting the assumptions generally yielded comparable or better classification accuracies than those employing the best single channel.

Methods for solving the general 2 class case were then proposed and programmed. No results are yet available. However, satisfactory results for this case are anticipated. In that event, a study of methods of attacking the general M class problem should be undertaken (see reference 2 of the report number 9.)

C. Probability density estimation and non-parametric classifiers

Several approaches have been taken in this area. G. de Montricher has proved the existence and uniqueness of the optimal (in the Bayes sense) probability density estimator. Unlike other schemes, this one is valid in any multidimensional space. Currently algorithms to find this function are being tested. However it is not yet clear how one would go about maximum likelihood classification of data in an economical manner using these functions. A report on the results of this work will be prepared by the end of August, 1973.

The approach in report number 2 involving spline smoothing of histograms is not readily extendable to multidimensional spaces. However, a new approach using cubic B-splines in Parzen's windows seems to offer a viable approach to non-parametric classification. Work is continuing on this and a report will be issued by August 31, 1973.

D. Use of spatial information in classification

Two approaches have been taken to employ spatial information to yield higher classification accuracies. One method is to smooth the raw data, thereby reducing non data-related variations. This method is discussed in report number 5. Also further research is presently being conducted and classification of the smoothed data will be performed. A report on these results will be available by August 31, 1973.

The second approach was to post-process the classified data points themselves. Here a simple scheme was devised that utilized the four nearest neighbors (4NN) of each point to verify and/or modify its classification. This algorithm resulted in significant improvement (~2-5%) in classification accuracy on two flight lines tested. Also it is very efficient. It appears particularly useful in classification of agricultural areas where fields contain many resolution elements.

Other algorithms for employing spatial information during classification time were also suggested. These and related schemes schemes should be investigated since much can be gained at apparently little cost in computer time.

E. Crop row reflectance model

This model enables one to compute the directional reflectance from a row crop. A row canopy model is used which takes edge effects into account. Only the development of the mathematical equations is contained in the report (Report No. 1).

F. Miscellaneous projects

In the course of the above projects, a few new techniques and interpretations have been reported, which, though not directly related to present problems at hand, could prove useful in these or related areas. These include an improved method for power spectral density estimation in the frequency domain, a random number generator for continuous random variables, and a geometrical interpretation of the 2n-th central difference. For completeness, these reports are included.

III. CONCLUSIONS

An administrative framework has been established, LARSYS in several forms is available, and a variety of projects have been undertaken during Phase 1 at the ICSA. Several projects are incomplete but shall be finished by the end of this summer, while research is continuing on others.

Final recommendations on these results include:

1) The Cholesky decomposition method should be employed in all programs (e. g. LARSYS, ERIPS) performing divergence and Gaussian maximum likelihood classification calculations
2) Research into employing spatial information in classification should be continued and encouraged

It is premature to make final recommendations on the other studies until their completion later this summer. Research is continuing in the feature extraction work, probability density estimation and non-parametric classifiers, and data smoothing techniques. Also methods (utilizing the Cholesky decomposition) for treating ill conditioned systems and accuracy studies are being investigated.

Appendix A

LARSYS AT ICSA

There are currently several versions of LARSYS available on the IBM 370/155 at Rice. Initially IBM FSD supplied ICSA with an OS version of LARSYS which they had written using the PS version supplied by LARS. The OS version reads LARSYS - 2 tapes but only has the statistics, classification, and display processors in LARSAA (no feature selection, per field classifier, or clustering).

Modifications to this version have been made and are currently installed. These include:

1) An overlaid LARSAA that executes in essentially the same amount of time and requires only 92K of memory (the original version requires 276K of memory).
2) A version of LARSAA that uses the Cholesky decomposition of the covariance matrix to perform the classification (see technical report number 7)
3) A version of LARSAA that employs the 4NN algorithm (uses spatial information) in classification (see technical report number 8)
4) A version of LARSAA that allows the user to employ linear combinations of channels as data rather than the individual channels (this provides a means for testing feature extraction results as in technical report number 9)
5) A version of LARSAA that uses only disk input and runs under the Time Sharing Option (TSO) enabling LARSAA to be run from a terminal.

All versions use the standard LARSYS input as documented in the "LARS RTCC User's Guide". Version 4 also uses some additional cards at the beginning describing the total number of channels, the number of linear combinations desired, and the coefficients for these combinations.

In addition to these versions, we have implemented a divergence calculation program for feature selection. This is used for timing and accuracy studies to be compared with a similar program that employs the Cholesky decomposition of the covariance matrices to perform the calculations.

Appendix B

LIST OF TECHNICAL REPORTS

1. " Mathematical Model Concerning Reflectance from a Row Crop" by R. K. Jaggi
2. "Spline Smoothing of Histograms by Linear Programming," by J. O. Bennett
3. "Power Spectral Density Estimation by Spline Smoothing in the Frequency Domain," by R. J. P. de Figueiredo and J. R. Thompson
4. "A Geometrical Interpretation of the 2n-th Central Difference," by R. A. Tapia
5. "Data Smoothing and Error Detection Based on Linear Interpolation," by V. M. Guerra and R. A. Tapia
6. "A Random Number Generator for Continuous Random Variables," by V. M. Guerra, R. A. Tapia and J. R. Thompson
7. "The Use of the Modified Cholesky Decomposition in Divergence and Classification Calculations, " by D. L. Van Rooy, M. S. Lynn, and C. H. Snyder
8. "Use of Spatial Information in Classification of Remotely Sensed Data" by D. L. Van Rooy and M. S. Lynn
9. "Optimal Feature Extraction - The Two Class Case" by W. S. Hsia and R. J. P. de Figueiredo